STAT 830
The Multivariate Normal Distribution

Richard Lockhart

Simon Fraser University

STAT 830 — Fall 2011

What I assume you already know

- The basics of normal distributions in 1 dimension.

The Multivariate Normal Distribution

- **Def'n**: $Z \in R^1 \sim N(0,1)$ iff

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} .$$

- **Def'n**: $Z \in R^p \sim MVN(0, I)$ if and only if $Z = (Z_1, \ldots, Z_p)^t$ with the Z_i independent and each $Z_i \sim N(0,1)$.
- In this case according to our theorem

$$\begin{aligned} f_Z(z_1, \ldots, z_p) &= \prod \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} \\ &= (2\pi)^{-p/2} \exp\{-z^t z/2\} ; \end{aligned}$$

  superscript t denotes matrix transpose.
- **Def'n**: $X \in R^p$ has a multivariate normal distribution if it has the same distribution as $AZ + \mu$ for some $\mu \in R^p$, some $p \times p$ matrix of constants A and $Z \sim MVN(0, I)$.

The Multivariate Normal Density

- Matrix A singular: X does not have a density.
- A invertible: derive multivariate normal density by change of variables:

$$X = AZ + \mu \Leftrightarrow Z = A^{-1}(X - \mu) \qquad \frac{\partial X}{\partial Z} = A \qquad \frac{\partial Z}{\partial X} = A^{-1}.$$

- So

$$\begin{aligned} f_X(x) &= f_Z(A^{-1}(x-\mu))|\det(A^{-1})| \\ &= \frac{\exp\{-(x-\mu)^t(A^{-1})^t A^{-1}(x-\mu)/2\}}{(2\pi)^{p/2}|\det A|}. \end{aligned}$$

The Multivariate Normal Density continued

- Now define $\Sigma = AA^t$ and notice that

$$\Sigma^{-1} = (A^t)^{-1}A^{-1} = (A^{-1})^t A^{-1}$$

and

$$\det \Sigma = \det A \det A^t = (\det A)^2 .$$

- Thus f_X is

$$\frac{\exp\{-(x-\mu)^t\Sigma^{-1}(x-\mu)/2\}}{(2\pi)^{p/2}(\det \Sigma)^{1/2}} ;$$

the $MVN(\mu, \Sigma)$ density.

- Note density is the same for all A such that $AA^t = \Sigma$.
- This justifies the notation $MVN(\mu, \Sigma)$.

The Multivariate Normal Density continued

- For which μ, Σ is this a density?
- Any μ but if $x \in R^p$ then, putting $y = A^t x$,

$$x^t \Sigma x = x^t A A^t x = (A^t x)^t (A^t x) = \sum_1^p y_i^2 \geq 0$$

- Inequality strict except for $y = 0$ which is equivalent to $x = 0$.
- Thus Σ is a positive definite symmetric matrix.
- Conversely, if Σ is a positive definite symmetric matrix then there is a square invertible matrix A such that $AA^t = \Sigma$ so that there is a $MVN(\mu, \Sigma)$ distribution.
- A can be found via the Cholesky decomposition, e.g.

Singular cases

- When A is singular X will not have a density.
- $\exists a$ such that $P(a^t X = a^t \mu) = 1$
- X is confined to a hyperplane.
- Still true: distribution of X depends only on $\Sigma = AA^t$
- if $AA^t = BB^t$ then $AZ + \mu$ and $BZ + \mu$ have the same distribution.
- Proof by mgfs in case $\mu = 0$:

$$\mathrm{E}(\exp(u^t AZ)) = \mathrm{E}(\exp(v^t Z))$$

where $v = A^t u$. Use the independent components of Z to get

$$\mathrm{E}(\exp(u^t AZ)) = \exp(v^t v/2) = \exp(u^t AA^t u) = \exp(u^t BB^t u) = \mathrm{E}(\exp(u$$

Properties of the *MVN* distribution

1. All margins are multivariate normal: if
$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$
$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$
and
$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$
then $X \sim MVN(\mu, \Sigma) \Rightarrow X_1 \sim MVN(\mu_1, \Sigma_{11})$.
2. All conditionals are normal: the conditional distribution of X_1 given $X_2 = x_2$ is $MVN(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$
3. $MX + \nu \sim MVN(M\mu + \nu, M\Sigma M^t)$: affine transformation of MVN is normal.